

Soft data, as outlined in this article, are present on the ground in plenty. We have endeavoured, using our growing skills, to do responsibly what we can. We have built a relationship of trust with our family health services authority. But trust alone will not be enough in the future, and we have to work to find methods of critically evaluating the benefits of the money spent. These evaluation techniques will need to look at the work that the medical audit advisory group has done and will need to find measures that accurately reflect the outcome of such work. It is unlikely that simply counting audits done in general practice will provide an adequate measure of the changes that have been brought about.

I thank Manchester general practitioners and their teams and the medical audit advisory group members and staff for making it possible to record so much progress.

- 1 Maynard A. Ivory tower: that's why the lady is a wimp. *Health Service Journal* 1993;103 (28 Jan):21.
- 2 Drummond MF. *Principles of economic appraisal in health care*. Oxford: Oxford University Press, 1980:116-9.
- 3 Baker R. Audit and standards in new general practice. *BMJ* 1991;303:32-4.

- 4 Grol R, Mokkink H, Schellevis F. The effects of peer review in general practice. *J R Coll Gen Pract* 1988;38:10-3.
- 5 Hughes J, Humphrey C. *Medical audit in general practice*. London: King's Fund, 1990.
- 6 Johnson R. Summarising GP records. *Audit Trends* 1993;1 (June):756-7.
- 7 Foxall M. Organising large-scale multi-practice audit projects. *Audit Trends* 1993;1 (September):99-101.
- 8 Benett IJ, Lambert C, Hinds G, Kirton C. Emerging standards for diabetes care from a city-wide primary care audit. *Diabetic Med* (in press).
- 9 Humphrey C, Berrow D. Developing role of medical audit advisory groups. *Quality in Health Care* 1993;2:232-8.
- 10 Department of Health. *Health service developments. Working for Patients: medical audit in the family practitioner services*. London: DoH, 1990 (HC(FP)90(8)).
- 11 House of Commons Health Committee. *Second report, 1991-92: maternity services*. London: HMSO, 1992. (Winterton Committee).
- 12 Department of Health. Expert Maternity Committee. *Changing Childbirth*. London: HMSO, 1993.
- 13 Royal College of General Practitioners. *What sort of doctor?* London: Royal College of General Practitioners, 1985.
- 14 Derry J, Lawrence M, Griew K, Anderson J, Humphrey J, Pandher K. Auditing audits: the method of Oxfordshire Medical Audit Advisory Group. *BMJ* 1991;303:1247-9.
- 15 Baker R. Use of psychometrics to develop a measure of patient satisfaction for general practice. In: Hopkins A, Fitzpatrick R, eds. *Measurement of patients' satisfaction with their care*. London: Royal College of Physicians, 1993.
- 16 General Medical Services Committee Audit Subcommittee. *The working MAAG*. London: BMA, 1991.

(Accepted 27 April 1994)

## Statistics Notes

### Diagnostic tests 2: predictive values

Douglas G Altman, J Martin Bland

This is the fourth in a series of occasional notes on medical statistics.

The whole point of a diagnostic test is to use it to make a diagnosis, so we need to know the probability that the test will give the correct diagnosis. The sensitivity and specificity<sup>1</sup> do not give us this information. Instead we must approach the data from the direction of the test results, using predictive values.

*Positive predictive value* is the proportion of patients with positive test results who are correctly diagnosed.

*Negative predictive value* is the proportion of patients with negative test results who are correctly diagnosed.

Using the same data as in the previous note,<sup>1</sup> we know that 231 of 263 patients with abnormal liver scans had abnormal pathology, giving the proportion of correct diagnoses as 231/263=0.88. Similarly, among the 81 patients with normal liver scans the proportion of correct diagnoses was 54/81=0.59. These proportions are of only limited validity, however. The predictive values of a test in clinical practice depend critically on the prevalence of the abnormality in the patients being tested; this may well differ from the prevalence in a published study assessing the usefulness of the test.

In the liver scan study the prevalence of abnormality was 0.75. If the same test was used in a different clinical setting where the prevalence of abnormality was 0.25 we would have a positive predictive value of 0.45 and a negative predictive value of 0.95. The rarer the abnormality the more sure we can be that a negative test indicates no abnormality, and the less sure that a positive result really indicates an abnormality. Predictive values observed in one study do not apply universally.

The positive and negative predictive values (PPV and NPV) can be calculated for any prevalence as follows:

$$\text{PPV} = \frac{\text{sensitivity} \times \text{prevalence}}{\text{sensitivity} \times \text{prevalence} + (1 - \text{specificity}) \times (1 - \text{prevalence})}$$

$$\text{NPV} = \frac{\text{specificity} \times (1 - \text{prevalence})}{(1 - \text{sensitivity}) \times \text{prevalence} + \text{specificity} \times (1 - \text{prevalence})}$$

If the prevalence of the disease is very low, the positive predictive value will not be close to 1 even if both the sensitivity and specificity are high. Thus in screening the general population it is inevitable that many people with positive test results will be false positives.

The prevalence can be interpreted as the probability, before the test is carried out that the subject has the disease, known as the prior probability of disease. The positive and negative predictive values are the revised estimates of the same probability for those subjects who are positive and negative on the test, and are known as posterior probabilities. The difference between the prior and posterior probabilities is one way of assessing the usefulness of the test.

For any test result we can compare the probability of getting that result if the patient truly had the condition of interest with the corresponding probability if he or she were healthy. The ratio of these probabilities is called the *likelihood ratio*, calculated as sensitivity/(1-specificity).

The likelihood ratio indicates the value of the test for increasing certainty about a positive diagnosis. For the liver scan data the prevalence of abnormal pathology was 0.75, so the pre-test odds of disease were 0.75/(1-0.75)=3.0. The sensitivity was 0.895 and the specificity was 0.628. The post-test odds of disease given a positive test is 0.878/(1-0.878)=7.22, and the likelihood ratio is 0.895/(1-0.628)=2.41. The post-test odds of having the disease is the pre-test odds multiplied by the likelihood ratio.

A high likelihood ratio may show that the test is useful, but it does not necessarily follow that a positive test is a good indicator of the presence of disease.

1 Altman DG, Bland JM. Diagnostic tests 1: sensitivity and specificity. *BMJ* 1994;309:0000000.

### Correction

#### Diet and cancer

A printer's error occurred in this article, the fourth in the series on cancer prevention in primary care by Joan Austoker (18 June, pp 1610-4). In table I the first mention of a possibly increased risk of cancer associated with increased fat intake (in the second column) should have applied to breast cancer and not to lung cancer as published.

Medical Statistics  
Laboratory, Imperial  
Cancer Research Fund,  
London WC2A 3PX  
Douglas G Altman, head

Department of Public  
Health Sciences, St  
George's Hospital Medical  
School, London  
SW17 0RE  
J Martin Bland, reader in  
medical statistics

BMJ 1994;309:102